



Analisis Kualitas Soal Asesmen Sumatif Akhir Semester Bahasa Jawa Kelas VII di SMPN 2 Kutowinangun

Anggilia Dwi Cahyani^{1*}, Joko Sukoyo²

anggiliadwiac@gmail.com^{1*}, j_sukoyo@mail.unnes.ac.id²

^{1,2}Program Studi Pendidikan Bahasa dan Sastra Jawa

^{1,2}Universitas Negeri Semarang

Received: 12 02 2025. Revised: 13 04 2025. Accepted: 22 05 2025.

Abstract : This research was aims to determine the quality of the Javanese final summative assessment questions for grade VII even semester at SMPN 2 Kutowinangun 2023/2024. This research used descriptive quantitative method with interview and study of document as data collection technique. The data source used in this research consisted of question sheets and 30 answer sheets wich taken randomly. The analysis of question was used SPSS dan anates. The results of this research showed (1) there are three questions should be revised because the construction is not complete; (2) based on validity, 57,14% questions was valid and 42,86% questions was invalid; (3) reliability of objective questions is very high, but the reliability of the essay questions is low; (4) difficulty level of the questions is not in balance proportion, so it needs to be revised; (5) based on descrimination level, 71,43% of questions was suitable; and (6) there found 9 questions with effective distractor, whereas 10 questions was not have effective distractor.

Keywords : Quality of Question, Summative Assessment, Javanese.

Abstrak : Penelitian ini bertujuan untuk mengetahui kualitas soal asesmen sumatif akhir semester genap bahasa Jawa kelas VII di SMPN 2 Kutowinangun tahun ajaran 2023/2024. Penelitian ini menggunakan metode deskriptif kuantitatif dengan teknik pengumpulan data melalui wawancara dan studi dokumen. Intrumen penelitian yang digunakan yaitu lembar soal dan 30 lembar jawab yang diambil secara acak. Analisis kualitas soal menggunakan SPSS dan *software anates*. Hasil penelitian menjelaskan bahwa (1) ditemukan tiga butir soal yang perlu direvisi karena tidak memenuhi konstruksi; (2) berdasarkan validitas, terdapat 57,14% soal valid dan 42,86% tidak valid; (3) reliabilitas soal objektif tergolong sangat tinggi, namun reliabilitas soal uraian tergolong rendah; (4) tingkat kesukaran soal belum menunjukkan proporsi yang seimbang sehingga perlu direvisi; (5) daya beda soal menunjukkan 71,43% soal layak digunakan; dan (6) ditemukan 9 butir soal soal dengan pengecoh yang efektif, sedangkan 10 soal belum memenuhi kriteria pengecoh yang efektif.

Kata Kunci : Kualitas soal, Asesmen sumatif, Bahasa Jawa.

PENDAHULUAN

Pembelajaran merupakan kegiatan yang tidak lepas dari istilah asesmen dan evaluasi. Asesmen berfungsi sebagai alat untuk mengukur pemahaman materi peserta didik secara individu (Meliana et al., 2020). Evaluasi berfungsi sebagai informasi yang dapat digunakan untuk menimbang suatu pengambilan keputusan. Asesmen dan evaluasi bisa digunakan untuk pengambilan keputusan. Perbedaan terletak pada tatarannya, asesmen mengacu pada informasi dari peserta didik (mikro), sedangkan evaluasi memiliki lingkup yang lebih luas seperti media, metode, asesmen, dll. Dapat disimpulkan bahwa asesmen termasuk bagian penting dalam evaluasi pembelajaran (Marzuki et al., 2023). Asesmen dilakukan secara berkala, baik awal, tengah maupun akhir pembelajaran. Dalam pembelajaran di Kurikulum Merdeka, asesmen dibagi menjadi asesmen diagnostik, asesmen formatif dan asesmen sumatif.

Asesmen diagnostik bertujuan mencari tahu kondisi awal peserta didik sehingga dapat merencanakan pembelajaran yang berdiferensiasi dan sesuai kebutuhan (Ardiansyah et al., 2023). Asesmen formatif bertujuan memonitori proses pembelajaran sehingga guru bisa mengamati kemajuan belajar peserta didik (Anggraena et al., 2022). Penjelasan ini sejalan dengan Permendikbudristek No. 21 Tahun 2022, asesmen formatif sebagai upaya pengumpulan informasi tentang perkembangan dan hambatan belajar peserta didik yang dapat menjadi *feedback* untuk perbaikan proses pembelajaran. Berbeda lagi dengan asesmen sumatif yang dilaksanakan pada akhir pembelajaran untuk memastikan tercapainya tujuan pembelajaran, baik satu tujuan ataupun beberapa tujuan sekaligus (Anggraena et al., 2022). Asesmen ini umumnya menjadi komponen perhitungan penilaian pada tengah semester, akhir semester, dan akhir jenjang.

Asesmen sumatif akhir semester (ASAS) adalah aktivitas yang berlangsung setiap akhir semester dan digunakan sebagai keputusan atas kinerja peserta didik. Hasil asesmen ini bisa menjadi bahan pertimbangan guru untuk meningkatkan mutu pembelajaran pada semester maupun tahun ajaran selanjutnya. Pelaksanaan asesmen sumatif akhir semester harus direncanakan secara optimal supaya dapat mengukur penguasaan peserta didik dan melihat ketercapaian tujuan pembelajaran selama satu semester. Salah satu instrumen terpenting dalam perencanaan asesmen sumatif akhir semester adalah butir soal. Penyusunan soal yang berkualitas sangatlah krusial. Tidak jarang guru menganggap pembuatan soal sebagai tantangan tersendiri. Soal yang berkualitas dapat memberikan informasi peserta didik secara akurat. Semakin baik soal yang dibuat, maka semakin baik pula fungsinya sebagai alat ukur. Guru perlu

mengadakan kegiatan analisis sebagai cara untuk memastikan kualitas soal (Fitriani, 2021; Maharani & Putro, 2020).

Analisis butir soal mencakup kegiatan pengumpulan, peringkasan, dan pemanfaatan informasi jawaban tes peserta didik untuk mengambil keputusan penilaian (Tanjung et al., 2024). Serangkaian analisis kualitas soal dilakukan setelah soal diujikan. Penyimpulan baik tidaknya kualitas soal bisa diketahui setelah menganalisis lima item yaitu validitas, reliabilitas, tingkat kesukaran, daya beda, dan keberfungsian pengecoh. Analisis butir soal dilakukan untuk mengategorisasi butir soal yang baik dan jelek. Bertujuan membantu guru untuk meningkatkan mutu soal dengan memilah soal yang perlu direvisi serta membuang soal tidak layak. Hasil analisis dapat digunakan sebagai acuan untuk membuktikan efektivitas soal sehingga dapat dimasukkan dalam bank soal yang berkualitas (Elgadal & Mariod, 2021). Analisis butir soal juga bermanfaat meningkatkan pembelajaran dan keterampilan guru terkait konstruksi soal.

Berdasarkan uraian di atas, proses analisis butir soal memiliki andil untuk meningkatkan keterampilan guru dalam membuat soal dan dapat menjadi bahan evaluasi untuk pembuatan soal ke depannya. Namun dalam praktiknya, banyak guru yang mengabaikan kegiatan analisis butir soal, seperti yang terjadi di SMP Negeri 2 Kutowinangun tahun ajaran 2023/2024. Berdasarkan hasil wawancara dengan guru bahasa Jawa kelas VII, guru yang ditugaskan wajib membuat 45 soal dengan rincian: 20 soal pilihan ganda, 10 soal benar-salah, 10 soal menjodohkan, dan 5 soal uraian. Selama bertugas sebagai guru bahasa Jawa sekaligus penyusun soal ASAS Kelas VII, guru belum pernah menganalisis kualitas soal karena belum menguasai caranya. Dengan demikian, soal asesmen yang dibuat belum diketahui kualitasnya sehingga peneliti tertarik untuk melakukan analisis guna mengetahui kualitas soal yang dibuat guru tersebut.

Penelitian sebelumnya yang relevan dengan penelitian ini adalah penelitian dari Maemunah (2015) yang menganalisis soal pilihan ganda dan uraian dalam ujian akhir semester ganjil dan genap Bahasa Jawa kelas VIII di SMP Negeri 2 Sruweng tahun pelajaran 2014/2015. Penelitian dengan metode deskriptif kuantitatif tersebut menggunakan *software Anates* dan *Microsoft Excel*. Berdasarkan hasil analisis, disimpulkan bahwa kualitas soal ujian akhir semester ganjil tergolong jelek dan kualitas soal ujian akhir semester genap tergolong cukup. Penelitian lainnya dari Zaki (2020) terhadap soal penilaian akhir semester dalam buku ajar bahasa Jawa kelas IV tahun pelajaran 2020/2021. Penelitian deskriptif kuantitatif ini berbantuan *software Anates version 4* dengan objek kajian 10 soal pilihan ganda dan 10 soal uraian. Berdasarkan hasil penelitian disimpulkan: 1) validitas soal pilihan ganda tergolong valid; 2)

reliabilitas soal pilihan ganda tergolong kurang; 3) daya pembeda soal pilihan ganda tergolong kurang; dan 5) efektivitas pengecoh soal pilihan ganda tergolong kurang. Kebaharuan penelitian ini dengan penelitian sebelumnya adalah peneliti menganalisis kualitas soal dengan bentuk yang lebih bervariasi, yaitu pilihan ganda, benar-salah, menjodohkan, dan uraian. Peneliti juga menggunakan *software SPSS* yang belum pernah digunakan dalam penelitian sebelumnya mengenai kualitas soal bahasa Jawa.

METODE PENELITIAN

Penelitian ini menggunakan metode deskriptif kuantitatif untuk mengetahui kualitas soal asesmen sumatif akhir semester (ASAS) genap bahasa Jawa Kelas VII di SMPN 2 Kutowinangun, Kabupaten Kebumen Tahun Ajaran 2023/2024. Kualitas soal dilihat berdasarkan validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh. Teknik pengumpulan data menggunakan wawancara dan studi dokumen. Teknik wawancara dilakukan dengan memberikan beberapa pertanyaan kepada guru bahasa Jawa sekaligus pembuat soal kelas VII seputar penyusunan soal. Teknik studi dokumen dilakukan dengan mengumpulkan lembar soal, lembar jawab, dan kunci jawaban. Sumber penelitian yaitu lembar soal ASAS dan 30 lembar jawaban peserta didik yang diambil secara acak. Data-data yang terkumpul kemudian dianalisis berbantuan SPSS dan *software Anates*. *Software SPSS* digunakan untuk menghitung validitas, reliabilitas, tingkat kesukaran dan daya beda soal pilihan ganda, benar-salah, dan menjodohkan. *Software Anates* digunakan dalam analisis terkait validitas, reliabilitas, tingkat kesukaran dan daya beda soal uraian, serta keefektifan pengecoh soal pilihan ganda.

HASIL DAN PEMBAHASAN

Pada lembar soal asesmen sumatif akhir semester (ASAS) genap Bahasa Jawa kelas VII di SMPN 2 Kutowinangun tahun ajaran 2023/2024, peneliti mendapati tiga butir soal yang tidak dapat diikutsertakan dalam penelitian. Butir soal pilihan ganda nomor 16 karena soal tidak lengkap, Butir soal menjodohkan nomor 33 dan 34 karena tidak ada pilihan jawaban benar. Pertanyaan dalam nomor soal tersebut tidak memenuhi konstruksi soal yang baik. Oleh karena itu, soal otomatis harus direvisi dan tidak memungkinkan untuk diteliti lebih lanjut. Validitas soal merupakan informasi mengenai sejauh mana sebuah butir soal memiliki kemampuan untuk mengukur capaian belajar peserta didik sesuai dengan kompetensi yang ingin diukur (Fatimah et al., 2020). Validitas soal dapat mengungkapkan relevansi suatu butir soal, apakah sudah

melingkupi aspek yang diinginkan, dan apakah jawaban dari pertanyaan mampu menggambarkan kondisi yang akan diukur (Tanjung et al., 2024).

Soal yang valid berdasarkan isinya mampu memberikan pengukuran tentang ketercapaian suatu capaian pembelajaran setelah proses penyampaian materi oleh guru. Sebuah butir soal dikatakan baik ketika memiliki validitas yang tinggi. Peneliti mengetahui valid atau tidaknya butir soal yang telah disusun guru dengan melihat nilai *sign (2-tailed)* dalam perhitungan *bivariate correlation* melalui *software SPSS*. Menurut (Sanusi & Aziez, 2021), soal yang valid memiliki nilai signifikansi lebih kecil dari 0,05 ($\text{Sign} < 0,05$). Berikut adalah hasil yang ditemukan berdasarkan *software SPSS* dan *Anates*:

Tabel 1. Validitas butir soal

Nomor Soal	Jumlah Soal	Persentase	Keterangan
4, 5, 6, 7, 10, 11, 14, 15, 17, 18, 19, 20, 22, 26, 28, 31, 32, 36, 37, 38, 40, 42, 44, 45	24	57,14%	Valid
1, 2, 3, 8, 9, 12, 13, 21, 23, 24, 25, 27, 29, 30, 35, 39, 41, 43	18	42,86%	Tidak Valid

Berdasarkan tabel 1, ditemukan butir soal berkategori valid sejumlah 24 (57,14%). Butir-butir soal tersebut dianggap mampu mengukur capaian pembelajaran bahasa Jawa peserta didik. Peneliti juga menemukan butir soal tidak valid sejumlah 18 (42,86%). Contoh soal valid yaitu pada butir soal nomor 26 dengan nilai signifikansi mencapai 0,001 yang berbunyi: “*Menawi bapak badhe kesah rumiyin mangga, kula dereng siram*”. Soal berbentuk pernyataan benar-salah ini disusun dengan tujuan untuk melihat sejauh mana capaian peserta didik mengenai penggunaan bahasa Jawa yang selaras dengan pedoman *unggah-ungguh basa*. Jawaban pernyataan dalam soal adalah salah (S) karena terdapat pemilihan kata yang tidak tepat yaitu *kesah* dan *siram*. Penggunaan diksi *kesah* merujuk pada diri sendiri dan orang yang lebih muda, sedangkan orang yang lebih tua atau orang yang dihormati seharusnya menggunakan diksi *tindak*. Begitu pula pada diksi *siram* (peruntukkan orang lebih tua/dihormati) harus diganti dengan *mandi*, sehingga kalimat yang benar “*Menawi bapak badhe tindak rumiyin mangga, kula dereng mandi*”. Hasil mengindikasikan soal dapat menjalankan fungsinya sebagai alat ukur yang absah.

Contoh soal tidak valid ada pada nomor 35 yang memperoleh nilai signifikansi sebesar 0,371. Soal tersebut berbunyi: “*Para Danawa bisa disingkirake. Danawa tegese ...*”. Soal bentuk menjodohkan ini memiliki tujuan untuk melihat perbendaharaan kata peserta didik melalui sinonim. Isi butir soal ini sebenarnya tidak berkaitan dengan capaian pembelajaran dalam kurikulum, sehingga dapat dikatakan bahwa soal tidak berfungsi sebagaimana mestinya

sebagai alat ukur yang absah. Berdasarkan temuan ini, guru perlu merevisi butir soal tidak valid terlebih presentase soal tidak valid memiliki selisih tidak terlalu jauh dari soal valid sehingga harus menjadi perhatian tersendiri bagi guru. Kualitas soal asesmen juga dapat dilihat dari besarnya nilai reliabilitas. Semakin besar nilai reliabilitas soal, maka semakin baik pula tingkat keajegannya dalam penilaian. Uji reliabilitas membantu guru untuk mengungkap sejauh mana soal dapat diandalkan (Fatimah et al., 2020). Bisa pula menggambarkan kestabilan soal jika dipakai berulang. Artinya ketika soal digunakan berulang dalam kesempatan berbeda akan menunjukkan hasil yang relatif sama (Farida & Musyarofah, 2021; Tarmizi et al., 2021).

Reliabilitas diklasifikasikan menjadi beberapa kategori berdasarkan nilai *cronbach's alpha*, antara lain: 0,00 – 0,20 (sangat rendah); 0,21 – 0,40 (rendah); 0,41 – 0,60 (cukup); 0,61 – 0,80 (tinggi); dan 0,81 – 1,00 (sangat tinggi). Menurut Heale & Twycross (dalam Anggraini et al., 2022), soal dinyatakan reliabel apabila nilai *cronbach's alpha* lebih besar dari 0,60 (> 0,60). Berikut adalah hasil reliabilitas soal asesmen berdasarkan perhitungan *software SPSS* dan *Anates*.

Tabel 2. Reliabilitas soal

<i>Cronbach's Alpha</i>	Bentuk Soal
0,869	Pilihan ganda, benar – salah, menjodohkan
0,40	Uraian

Interpretasi tabel 2 menunjukkan bahwa soal pilihan ganda, benar – salah, dan menjodohkan tergolong reliabel dengan nilai sebesar 0,869 (reliabilitas sangat tinggi). Perhitungan ini menunjukkan jika soal dapat dipercaya, dapat memberikan hasil asesmen yang konsisten, dan dapat dipergunakan dengan peserta didik yang berbeda. Berdasarkan nilai *cronbach's alpha*, soal-soal dapat dimanfaatkan sebagai alat ukur untuk asesmen yang akan datang ataupun dimasukkan dalam bank soal. Soal uraian termasuk soal tidak reliabel karena nilai *cronbach's alpha* kurang dari batas minimal, yaitu hanya 0,40 (reliabilitas rendah) dan mengindikasikan apabila soal digunakan berulang kali maka berpotensi memberikan hasil yang berbeda. Dengan demikian, guru perlu merevisi soal uraian.

Selain validitas dan reliabilitas, kualitas soal dapat ditunjang dengan adanya proporsi tingkat kesukaran yang ideal. Tingkat kesukaran merupakan informasi untuk menggambarkan seberapa sukar soal asesmen. Asumsi kesukaran soal harus berdasarkan sudut pandang peserta didik dengan menitikberatkan pada kesanggupan mereka dalam menjawab, bukan menurut pandangan guru semata (Sanusi & Aziez, 2021). Semakin banyak peserta didik menjawab benar

maka sebuah butir soal dikatakan mudah, begitupun sebaliknya. Tingkat kesukaran soal diklasifikasikan menjadi tiga kategori, yaitu mudah, sedang, dan sukar.

Penyusunan butir soal asesmen sebaiknya jangan cenderung mudah ataupun sukar. Soal dengan tingkat kesulitan mudah tidak akan mampu mengukur pemahaman peserta terhadap suatu materi secara mendalam, namun soal yang sulit belum tentu termasuk soal yang baik. Soal yang mudah tidak dapat menstimulasi peserta didik untuk berusaha menjawab. Lain hal soal yang terlalu sulit berpotensi mematahkan semangat menjawab karena mereka merasa sesuatu yang ditanyakan berada di luar kemampuannya. Bukan berarti soal kategori mudah dan sulit tidak boleh digunakan dalam asesmen. Justru keberadaan soal tersebut menjadi variasi tersendiri selama masih sesuai proporsinya. Rasio tingkat kesukaran yang ideal dalam soal asesmen yaitu 25% soal mudah, 50% soal sedang, dan 25% soal sukar (Bano et al., 2022). Analisis tingkat kesukaran soal melalui *software SPSS* dilakukan dengan melihat nilai *mean* butir soal. Nilai yang dihasilkan kemudian diinterpretasikan dalam beberapa kategori, antara lain: “mudah” dengan interval nilai 0,71 – 1,00; “sedang” dengan interval nilai 0,31 – 0,70; dan “sukar” dengan rentang interval 0,00 – 0,30 (Kusuma, 2024). Hasilnya tertera dalam tabel berikut.

Tabel 3. Tingkat kesukaran butir soal

Nomor Soal	Jumlah Soal	Persentase	Kategori
1, 3, 6, 9, 10, 15, 21, 22, 23, 24, 27, 29, 30, 31, 32, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45	26	61,90%	Mudah
2, 4, 5, 7, 8, 11, 12, 14, 17, 18, 19, 20, 25, 26, 28	18	35,70%	Sedang
13	1	2,4%	Sukar

Data dalam tabel 3, menunjukkan 61,90% soal berkategori “mudah”; 35,70% berkategori “sedang”; dan 2,4% berkategori “sukar”. Hasil menggambarkan belum adanya proporsi tingkat kesukaran yang seimbang. Sebagian besar soal yang dibuat guru masih tergolong mudah, sedangkan soal sulit hanya ditemukan satu butir. Berdasarkan hasil, guru perlu merevisi beberapa soal supaya mendapat keseimbangan yaitu mengurangi jumlah soal mudah dan menambahkan soal sedang/sulit. Berikut disajikan contoh soal berdasarkan kategorinya.

22. Ing Dieng hawane panas banget

[...]

Gambar 1. Contoh soal kategori mudah

14. Bharata nyuwun supados Rama Wijaya kondur Ayodya. Ukara kasebut menawa diowahi dadi ukara tanggap yaiku
- A. Bharata disuwun supados Rama Wijaya kondur Ayodya
 - B. Rama Wijaya nyuwun supados Bharata kondur Ayodya
 - C. Rama Wijaya dipunsuwun Bharata supados kondur Ayodya
 - D. Bharata dipunsuwun supados Rama Wijaya kondur Ayodya

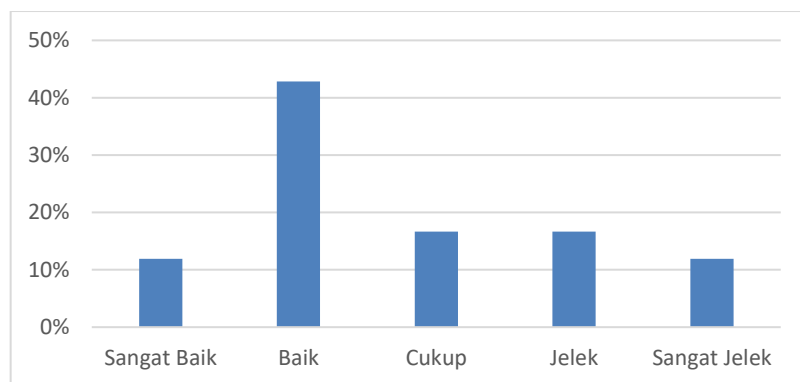
Gambar 2. Contoh soal kategori Sedang

13. Prabu Dasaratha tundane seda jalaran ora kuwawa nandhang duhkita. Dhuhkitane Prabu Dasarata yaiku
A. Raden Bharata ora gelem dadi raja
B. Pisah karo Rama Wijaya amarga kalah perjanjian
C. Dewi Kekayi njaluk supaya janjine diugemi
D. Ora lila menawa kudu seleh keprabon

Gambar 3. Contoh soal kategori sukar

Aspek lain yang harus dianalisis untuk mengetahui kualitas soal adalah daya beda soal. Daya beda soal berkaitan dengan keberfungsian soal untuk membedakan peserta didik kelompok atas dan kelompok bawah. Artinya, melalui analisis daya pembeda guru dapat menentukan peserta didik mana yang pandai dan tidak pandai (Fiska et al., 2021). Anggapannya ketika soal berdaya beda baik diberikan, maka peserta didik berkemampuan tinggi akan mampu mengerjakan dan seharusnya tidak mampu dikerjakan oleh peserta didik berkemampuan rendah. Keadaan demikian dapat terjadi karena persiapan pra asesmen yang tentunya berbeda di antara keduanya.

Daya beda dapat diketahui dengan melihat nilai *correlation product moment* (r) pada *software SPSS* yang kemudian dibandingkan dengan kategori berikut: “sangat jelek” dengan rentang nilai $< 0,00$; “jelek” dengan rentang nilai $0,00 - 0,20$; “cukup” dengan rentang nilai $0,21 - 0,40$; “baik” dengan rentang nilai $0,41 - 0,70$; dan “sangat baik” dengan rentang nilai $> 0,70$. Butir soal yang layak digunakan dalam asesmen yaitu soal yang minimal berada pada kategori “cukup” karena kemungkinan perbaikan tentu akan lebih sedikit dibandingkan yang berkategori “jelek” dan “sangat jelek” (Sari et al., 2022). Hasil analisis tercantum dalam gambar berikut ini.

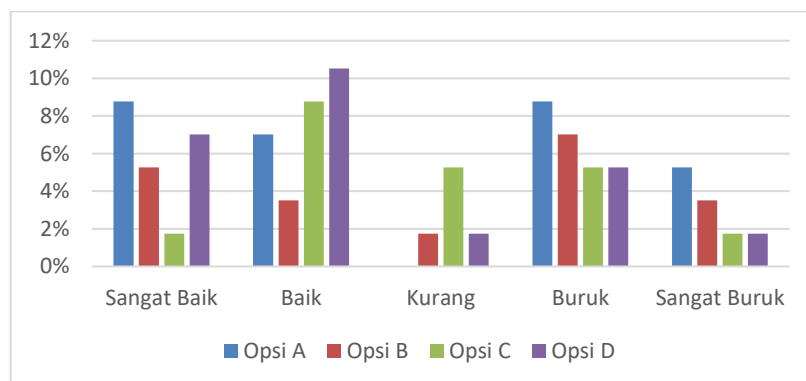


Gambar 4. Persentase daya beda butir soal

Berdasarkan gambar 3, dapat diketahui persentase yang diperoleh setiap kategori yaitu 11,90% butir soal mendapat kategori “sangat baik”; 42,86% butir soal mendapat kategori “baik”; 16,67% butir soal mendapat kategori “cukup”; 16,67% butir soal mendapat kategori “jelek”; dan 11,90% butir soal mendapat kategori “sangat jelek”. Lebih jelasnya soal dengan kategori “sangat baik” ditemukan sejumlah 5 butir; soal kategori “baik” sejumlah 18 butir; soal

kategori “cukup” sejumlah 7 butir; soal kategori “jelek” sejumlah 7 butir; dan soal kategori “sangat jelek” sejumlah 5 butir. Berdasarkan data, sebagian besar butir soal layak untuk asesmen dan hanya terdapat 12 butir soal yang tidak layak.

Identifikasi kualitas soal selanjutnya yaitu efektivitas pengecoh. Pengecoh merupakan pilihan jawaban salah yang bertujuan untuk mengelabui peserta didik dalam mengerjakan soal pilihan ganda. Pengecoh yang efektif harus memenuhi beberapa kriteria dalam pembuatannya, antara lain: dibuat homogen dan logis sesuai dengan kunci jawaban; rumusan pengecoh ditulis dengan panjang yang relatif sama dengan kunci jawaban; tidak memperlihatkan perbedaan yang terlalu mencolok; dan rumusan tidak boleh menyatakan bahwa semua opsi benar. Pengecoh yang efektif berfungsi apabila dipilih minimal 5% dari seluruh populasi (Sari et al., 2022). Peserta didik yang terkecoh sebaiknya lebih cenderung berasal dari kelompok kurang pandai walaupun faktanya tetap ada kelompok pandai yang tidak luput memilih pengecoh karena kurang memahami materi tertentu. Peneliti menggunakan *software Anates* dalam menganalisis keefektifan pengecoh. Opsi jawaban diinterpretasikan menjadi: sangat buruk, buruk, kurang, baik, dan sangat baik. Berdasarkan hasil analisis dari 57 pengecoh yang ada dalam soal, berikut distribusi setiap kategorinya.



Gambar 5. Persentase distribusi pengecoh

Berdasarkan gambar 3, diketahui persentase efektivitas pengecoh pada masing-masing kategori. Hasil menunjukkan dari 57 pengecoh dalam soal, kategori “sangat baik” sebanyak 22,80% atau 13 pengecoh. Kategori “baik” sebanyak 29,82% atau 17 pengecoh. Kategori “kurang” sebanyak 8,77% atau 5 pengecoh. Persentase kategori “buruk” sebesar 26,32% atau 15 pengecoh, dan 12,29% atau 7 pengecoh termasuk kategori “sangat buruk”. Dari 19 butir soal pilihan ganda yang diikutsertakan dalam analisis keefektifan pengecoh dapat disimpulkan bahwa hanya ada 9 yang layak untuk diujikan, yaitu soal nomor 2, 5, 6, 8, 11, 13, 14, 17, dan 19. Kelayakan soal-soal tersebut dikarenakan setiap opsi pengecoh berfungsi dan dipilih oleh paling sedikit satu peserta didik. 10 soal lainnya belum memenuhi kriteria pengecoh yang

efektif karena ada opsi jawaban yang sama sekali tidak memberikan daya tarik bagi peserta didik untuk memilihnya.

SIMPULAN

Berdasarkan hasil analisis, peneliti menyimpulkan bahwa kualitas soal asesmen sumatif akhir semester genap bahasa Jawa kelas VII di SMPN 2 Kutowinangun tahun ajaran 2023/2024 sebagai berikut: (1) terdapat tiga butir soal yang harus direvisi karena konstruksi soal tidak lengkap sehingga harus diperbaiki. (2) Terdapat 24 butir soal valid dengan persentase 57,14% dan 18 soal tidak valid dengan persentase 42,86%. (3) Reliabilitas soal objektif tergolong sangat tinggi dengan nilai $r = 0,869$ dan reliabilitas soal uraian tergolong rendah dengan nilai $r = 0,40$. (4) Tingkat kesukaran soal belum menunjukkan proporsi yang seimbang, yaitu 61,90% soal berkategori “mudah”, 35,70% soal kategori “sedang”, dan 2,4% soal kategori “sukar”. (5) Daya beda soal menunjukkan 71,43% soal layak digunakan dengan rincian 11,90% soal kategori “sangat baik”, 42,86% soal kategori “baik”, 16,67% soal kategori “cukup”, 16,67% soal kategori “jelek”, dan 11,90% soal kategori “sangat jelek”. (6) Ditemukan 9 butir soal dengan pengecoh yang efektif, sedangkan 10 soal belum memenuhi kriteria pengecoh yang efektif. Berdasarkan data-data di atas, dapat disimpulkan bahwa kualitas soal asesmen sumatif akhir semester genap bahasa Jawa kelas VII di SMPN 2 Kutowinangun tahun ajaran 2023/2024 tergolong cukup. Penarikan simpulan ini diambil karena hasil persentase pada setiap item yang dianalisis rata-rata memiliki selisih yang tidak terlalu jauh antara soal layak dan yang tidak layak, selain itu penemuan 3 soal yang tidak layak secara konstruksi turut menjadi bahan pertimbangan. Dengan demikian, guru perlu merevisi beberapa soal terutama terkait dengan validitas, tingkat kesukaran, dan pengecoh supaya menghasilkan soal dengan kualitas lebih baik ke depannya. Melalui penelitian ini diharapkan dapat membantu guru untuk meningkatkan kompetensi terkait penyusunan soal asesmen bahasa Jawa.

DAFTAR RUJUKAN

Anggraena, Y., Ginanto, D., Felicia, N., Andiarti, A., Herutami, I., Alhapip, L., Iswoyo, S., Hartini, Y., & Mahardika, R. L. (2022). *Panduan Pembelajaran dan Asesmen Pendidikan Anak Usia Dini, Pendidikan Dasar, dan Menengah*. Badan Standar, Kurikulum, dan Asesmen Pendidikan Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia. <https://kurikulum.kemdikbud.go.id/wp-content/uploads/2022/06/Panduan-Pembelajaran-dan-Asesmen.pdf>

- Anggraini, F. D. P., Aprianti, A., Setyawati, V. A. V., & Hartanto, A. A. (2022). Pembelajaran Statistika Menggunakan Software SPSS untuk Uji Validitas dan Reliabilitas. *Jurnal Basicedu*, 6(4), 6491–6504. <https://doi.org/10.31004/basicedu.v6i4.3206>
- Ardiansyah, Sagita, F., & Juanda. (2023). Assesmen dalam Kurikulum Merdeka Belajar. *Jurnal Literasi Dan Pembelajaran Indonesia*, 3(1), 8–13. <https://jurnal.fkip.samawa-university.ac.id/JLPI/article/view/361>
- Bano, V. O., Marambaawang, D. N., & Njoeroemana, Y. (2022). Analisis Kriteria Butir Soal Ujian Sekolah Mata Pelajaran IPA di SMP Negeri 1 Waingapu. *Ideas: Jurnal Pendidikan, Sosial, Dan Budaya*, 8(1), 145. <https://doi.org/10.32884/ideas.v8i1.660>
- Elgadal, A. H., & Mariod, A. A. (2021). Item analysis of multiple-choice questions (MCQs): assessment tool for quality assurance measures. *Sudan Journal of Medical Sciences*, 16(3), 334–346. <https://www.ajol.info/index.php/sjms/article/view/216888>
- Farida, & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab*, 1(1), 34–44. <https://doi.org/10.32923/al-muarrib.v1i1.2100>
- Fatimah, S., Elzamzami, A. B., & Slamet, J. (2020). Item Analysis of Final Test for the 9th Grade Students of SMPN 44 Surabaya in the Academic Year of 2019/2020. *JournEEL (Journal of English Education and Literature)*, 2(1), 34–46. <https://doi.org/10.51836/journeel.v2i1.81>
- Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis Butir Soal Ulangan Harian IPA Menggunakan Software Anates pada Pendekatan Teori Tes Klasik. *Natural Science Education Research*, 4(1), 65–76. <https://doi.org/10.21107/nser.v4i1.8133>
- Fitriani, N. (2021). Analisis Tingkat Kesukaran, Daya Pembeda, dan Efektivitas Pengecoh Soal Pelatihan Kewaspadaan Kegawatdaruratan Maternal dan Neonatal. *Paedagoria : Jurnal Kajian, Penelitian Dan Pengembangan Kependidikan*, 12(2), 199. <https://doi.org/10.31764/paedagoria.v12i2.4956>
- Kusuma, W. B. (2024). Validitas dan Reliabilitas Materi Uji Kompetensi Ruang Lingkup Keselamatan dan Kesehatan Kerja LSP PPSDM Migas Tahun 2023. *Jurnal Nasional Pengelolaan Energi MigasZoom*, 6(1), 25–32. <https://doi.org/10.37525/mz/2024-1/561>
- Maemunah, S. (2015). Analisis Butir Soal Ujian Akhir Semester Mata Pelajaran Bahasa Jawa Kelas VIII SMP Negeri 2 Sruweng Kebumen Tahun Pelajaran 2014/2015. *Jurnal Program Studi Pendidikan Bahasa Dan Sastra Jawa Universitas Muhammadiyah*

Purworejo, 7(08), 101–108.
<http://ejournal.umpwr.ac.id/index.php/aditya/article/view/2435>

- Maharani, A. V., & Putro, N. H. P. S. (2020). Item Analysis of English Final Semester Test. *Indonesian Journal of EFL and Linguistics*, 5(2), 491. <https://doi.org/10.21462/ijefl.v5i2.302>
- Marzuki, I., Sholihah, T., & Imansyah, F. A. (2023). Urgensi Aspek Penilaian dalam Evaluasi Pembelajaran. *Tadarus Tarbawy: Jurnal Kajian Islam Dan Pendidikan*, 5(1), 1–6. <https://doi.org/10.31000/jkip.v5i1.8634>
- Meliana, S., Sudi Utami, E., & Sukoyo, J. (2020). Pengembangan Penilaian Autentik Pembelajaran Teks Dialog Bahasa Jawa Kelas VII. *Piwulang: Jurnal Pendidikan Bahasa Jawa*, 8(1), 13–17. <https://doi.org/10.15294/piwulang.v8i1.29581>
- Sanusi, R. N. A., & Aziez, F. (2021). Analisis Butir Soal Tes Objektif dan Subjektif untuk Keterampilan Membaca Pemahaman pada Kelas VII SMP N 3 Kalibagor. *Metafora: Jurnal Pembelajaran Bahasa Dan Sastra*, 8(1), 99. <https://doi.org/10.30595/mtf.v8i1.8501>
- Sari, V. N. I., Utomo, A. P. Y., & Sumarwati. (2022). Kualitas Soal Bahasa Indonesia di SMP Muhammadiyah 1 Pontianak: Analisis Butir Soal. *Jurnal Pendidikan Dan Sastra Indonesia*, 11(2), 112–119. <https://doi.org/10.15294/jpbsi.v11i2.58091>
- Tanjung, M. A. H. R., Fahmi, A. A., Rahmanita, F., Habibah, I. F., & Qomari, N. (2024). Analisis Butir Soal Penilaian Akhir Tahun Pelajaran Bahasa Arab Kelas VII MTs Al-Ma'arif Rakit Banjarnegara Jawa Tengah. *Mantiqutayr: Journal of Arabic Language*, 4(1), 347–367. <https://doi.org/10.25217/mantiqutayr.v4i1.4038>
- Zaki, M. (2020). Analisis Kualitas Butir Soal Bahasa Jawa untuk Penilaian Akhir Semester Madrasah Ibtida'iyah Tahun Pelajaran 2020/2021. UIN Sunan Kalijaga Yogyakarta. <https://digilib.uin-suka.ac.id/id/eprint/44788>